Chapter 17
Soybean Improvement and the Role of Gene Editing

Nihal Öztolan Erol

Abstract Soybean is a major agricultural crop that is used for food, feed, and industrial products. However, soybean production is facing several challenges, including pests, diseases, and environmental factors. In recent years, there has been a growing interest in using gene editing technologies to improve soybean traits. Gene editing technologies offer a promising new approach to improving soybean production and quality.

Gene editing technologies can be used to precisely alter the soybean genome. There are a number of different gene editing technologies that can be used to improve soybeans. One of the most commonly used technologies is CRISPR/Cas9, which uses a protein called Cas9 to cut DNA at a specific location. This can be used to insert, delete, or modify genes. Other gene editing technologies include zinc finger nucleases (ZFNs), and transcription activator-like effector nucleases (TALENs). Gene editing technologies have the potential to revolutionize soybean breeding. This can be used to introduce new traits, such as resistance to pests and diseases, or to improve existing traits, such as yield and oil content.

The use of gene editing technologies in soybean improvement is still in its early stages, but the potential benefits are significant. Gene editing technologies offer a more precise and efficient way to improve soybean production than traditional breeding methods. They also offer the potential to create new varieties of soybeans that are better able to meet the challenges of a changing world.

1 Soybean Production and Its Economic Value

Soybean is a very important crop that provides substantial oil and protein nutrition for the increasing human population. Soybean cultivation has been rooted back in ancient times c. 6000–9000 years ago, in East Asia [1]. Its massive production has reached its highest in the last century with the help of improving breeding

N. Ö. Erol (✉)
Sabancı University Nanotechnology Research and Application Center, İstanbul, Turkey
e-mail: nihal.oztolan@sabanciuniv.edu

A. Ricroch et al. (eds.), *A Roadmap for Plant Genome Editing*,
https://doi.org/10.1007/978-3-031-46150-7_17

techniques. Soybean production has increased since 1961 from 20–30 million tons to 350 million tons per year [2].

Soybean is very rich in oil and protein compared to other members of the legume family. Therefore, it meets a considerable demand for animal feed and oil production. Over three-fourths of soybean by weight is used for feeding livestock, poultry, and aquaculture production, so some countries increasingly export soybean products while others import to meet the demand for soybean-based animal feed. The rest is consumed by humans as an industrial oil, biofuel, food ingredients (lecithin, emulsifier, and proteins), and food (soy sauce, tempeh, soy milk, and tofu). Soybean is introduced as a rich protein source for plant-based diets as it consists of 40% of the dry matter including nine essential amino acids. Therefore, it is very important for the vegetarian and vegan diet, it provides high nutrition with its protein content [3]. Soybean seeds are the most important part of the plant, so throughout the domestication process, traits improving soybean seed quality and yield have been artificially selected for efficient utilization in the food industry and agriculture.

Soybean domestication has led to a significant reduction in genetic diversity due to selective sweeps, resulting in the fixation of beneficial traits. Studies have shown that nucleotide fixation during soybean domestication and improvement has resulted in a reduction of genetic diversity compared to wild soybean populations. Furthermore, the fixation of key genes involved in the regulation of traits such as seed size, pod dehiscence, and photoperiodic flowering has played a crucial role in shaping the morphology and adaptation of soybean to different environments. These genetic changes have contributed to increased yield and better adaptation to a range of environmental conditions, making soybean a globally important crop. However, the reduced genetic diversity resulting from selective sweeps also raises concerns regarding the resilience and adaptability of soybean crops in the face of new and changing environmental challenges.

2 Genetically Modified Soybean

Genetic modification of an organism traces back to the domestication of organisms. However, public perception misinterprets this term and people think that genetic modification of organisms came out with the developments in biotechnology in the late twentieth century. Breeding practices have long been used by humans and are striking evidence of genetic modification. With the discovery of recombination techniques in bacteria, genetic modification techniques have been gradually improved over the years and used first in producing medicines and then crops.

Recombinant DNA technologies are the fundamentals of genetic modifications in living organisms. Briefly, a target gene cassette-containing vector is transferred by a virus or a bacterium into living cells of an organism to insert a specific genetic sequence into that organism's genome. The first genetically modified soybean was produced in the 1990s. Glyphosate-resistant soybean cultivation together with the glyphosate herbicide dramatically decreased labour that occurred due to tillage of

the soil, at the same time this dual application increased genetically modified soybean production.

Despite its recalcitrant nature of regeneration under tissue culture techniques, several studies showed that soybean has been used for gene editing of flowering time, seed oil content, lateral root growth, and defence mechanism [4] (Table 17.1).

Curtin et al. (2011) first published the research about hairy-root and whole-plant transformation mediated by *Agrobacterium rhizogenes* using the zinc-finger nuclease (ZFN) method to target *DICER LIKE (DCL), RNA-DEPENDENT RNA POLYMERASE (RDR),* and *HUA ENHANCER1 (HEN1)* genes in the root cells [5].

Table 17.1 Soybean traits and associated genes that are modified by gene editing techniques CRISPR/Cas9, zinc-finger nuclease (ZFN), transcription activator-like effector nucleases (TALEN), *Agrobacterium tumafaciens,* and bean pod mottle virus mediated transformations

Trait	Gene	Gene editing technique	References
Pod shattering	*GmPdh1*	CRISPR/Cas9	[27]
Determinacy	*GmTfl1*		[33]
	GmSOC1		[35]
	GmLHY		[32]
Photoperiodicity	*GmFT2*		[39]
	GmE2		[42]
	GmPRR37		[15]
	GmFKF1		[37]
	GmCPDK38		[43]
Photomorphogenesis and flowering time	*GmPHYA/ GmPHYB*		[38]
Seed sugar transportation, seed oil and protein content	*GmSWEET10*		[15]
Seed sugar transportation, seed oil and protein content Stachyose content	*GmMFT*		[48]
	GmSTS		[17]
Rafinose content	*GmRS*		[16]/ [17]
Lipoxigenase	*GmLox*		[15]
Salt stress tolerance	*GmAITR*		[66]
Seed size	*GmSSS1*		[54]
Seed thickness	*GmST1*		[74]
	GmDCL		[5]/ [6]
	GmRDR	ZFN ZFN/TALEN	[5]
	GmHEN1		[5]
Fatty acid	*GmFAD2–1*		[7]/ [8]
Fatty acid	*GmFAD3*	TALEN	[9]
Albino	*GmPDS*	TALEN	[10]
Drought resistance	*GmNFYA13*	*Agrobacterium tumefaciens*	[64]
Seed hardness	*GmHs1*	*Agrobacterium tumefaciens* Bean pod mottle virus	[44]
Seed size, biotic and abiotic stress resistance	*GmFAD3*		[53]

They continued using the ZFN method to create double mutants of *DCL1a* and *DCL1b* for *Agrobacterium rhizogenes*-mediated whole-plant transformation [6]. ZFNs were also used to deliver multiple different DNA donors in *the FAD2-1a* locus (Glyma.10 g278000) by using a biolistic bombardment technique on immature embryo explants [7]. This research successfully regenerated fertile plants and transmitted the insert to the next generation.

FAD2-1a and *FAD2-1b* loci were mutated by using transcription activator-like effector nucleases (TALEN), to convert oleic acid into linoleic acid to increase the polyunsaturated fatty acid. The study used the hairy-root transformation method mediated by *Agrobacterium rhizogenes* [8]. The same research group also targeted *FAD2-1a, FAD2-1b,* and *FAD3* to convert linoleic acid into oleic acid by using the TALEN technique and again successfully transformed soybean immature embryo explants [9].

Du et al. (2016) conducted a study to compare two gene editing techniques, TALEN and CRISPR-Cas9 in parallel with testing the transformation efficiency by using soybean-specific *U6-10* and *Arabidopsis*-specific *U6-26* promoters in soybean [10]. They targeted a gene encoding phytoene desaturase (PDS), a rate-limiting enzyme involved in the carotenoid biosynthesis pathway. Hairy root transformation mediated by *Agrobacterium rhizogenes* successfully resulted in mutated buds. The study suggested the usage of CRISPR/Cas9 with species-specific promoters to acquire a highly transformation-efficient, cost-efficient, and easy-to-construct transformation technique [10].

CRISPR/Cas9 technique outperforms other precise gene editing techniques by its cost-efficient and easily applicable features. This technique accelerates soybean breeding and supports soybean production. Cytoplasmic male sterility (*GmAMS1*) [11], flowering time (*LNK2*) [12], seed oil content (*GmFAD2*) [13, 14]), seed lipoxygenase, raffinose, and stachyose contents (*GmLox, GmRS, GmSTS*) [15–17], plant growth and structure (*GmLHY* and *GmSPL9*) [18, 19] are some of the traits, which have a significant role in soybean breeding, that were studied in the last decade.

Investigation of causal alleles of certain traits has provided the most crucial information for gene editing applications. QTL mapping, using GWAS and linkage mapping analysis, along with functional genetic mutations unravel the causative nucleotide changes. With the introduction of nucleic acid-cutting enzymes and nucleases, gene editing breathes new life into plant breeding by precise editing, and soybean breeding will definitely benefit from this transformative new breeding techniques.

3 Agronomically Important Soybean Traits and the Use of Gene Editing

Soybean breeding plays a crucial role in the production of soybeans around the world as it helps to develop soybean varieties that can adapt to different environmental conditions. By breeding soybean varieties that are good quality, and tolerant to various biotic and abiotic stresses, soybean production can be increased and stabilized. Additionally, breeding efforts have resulted in soybean varieties with

desirable traits such as high yield, improved nutritional quality, and enhanced oil and protein content, which are important for meeting the increasing demand for soybean products worldwide. Overall, soybean breeding has been instrumental in improving soybean production by developing varieties that are better adapted to the diverse environmental conditions in different regions of the world. A rapid and precise gene editing might help to improve elite soybean varieties. Soybean improvement might be accelerated by introducing a non-synonymous mutation with the help of gene editing.

Although trading and migrating routes had caused the dissemination of a certain type of cultivated soybean seeds towards Eastern Asia and North America, local landraces had provided efficient genetic resources for soybean breeding in adaptation to the environment. The idea of a single origin of soybean domestication does not completely explain the existence of allelic variation among cultivated varieties. Because local genetic diversity had provided location-specific causal alleles associated with the traits of interest. Therefore, to improve plants' yield capacity and seed quality in terms of oil and protein, there were several genes functionally identified to be responsible for plant architectural, physiological, and morphological changes in organs by using CRISPR/Cas9 method.

3.1 Pod Shattering Resistance

Pod shattering resistance, to prevent seed dispersal and yield loss, is an important agronomical trait that has come along with domestication [20]. Angiosperms develop their seeds within the fruit and disperse them when there is an abscission between pedicel and lemma. This decreases the harvest output and was taken under control by artificially selecting pod-shattering resistant plants. Four pod-shattering resistance-associated genes were identified in soybean: *GmSHAT1-5, Pdh1, NST1A,* and *Glyma09g06290*.

Dong et al., (2014) identified a causal polymorphism in *the GmSHAT1-5* gene and the pod-shattering resistant domesticated soybeans, which were diversified from wild soybeans, derived from this single haplotype [21]. *GmSHAT1-5* is responsible for the lignification of fiber cap cells in the pod ventral suture which causes thickening in domesticated soybeans. The sample collection included both *Glycine max* and *Glycine soja* varieties gathered from the seed bank of the Chinese Academy of Agricultural Sciences (Beijing). The pod indehiscent allele from *Glycine max* showed a 13-fold higher expression than *Glycine soja*. It seems like domestication significantly affects pod-shattering traits; however, this research did not reveal the origin of the indehiscent allele.

Zhang and Singh (2020) identified a locus called *NST1A,* which showed epistasis with *Pdh1*. *NST1A* was a NAC family gene, a paralog of *GmSHAT1-5* [22]. Likewise, in NAC family transcription factors in *Arabidopsis thaliana*, a premature stop codon was identified to be responsible for gain-of-function mutation, where it provided pod shattering-resistance despite the existence of *the Pdh1* allele [22, 23]. The indehiscent *NST1A* allele was predominantly found in Southern China and Japan, this

implies that local wild cultivars in those regions were selected for the indehiscent *NST1A* allele independent of low humid conditions.

A genome-wide association study genotyped 211 soybean accessions including modern and wild cultivars collected from the National Center for Soybean Improvement of China by using NJAU 355 K SoySNP array containing 282,469 SNPs. A quantitative trait locus was identified on chromosome nine and within that locus, a candidate gene *Glyma09g06290* was found homologous to *Arabidopsis thaliana* basic helix-loop-helix, a gene responsible for silique dehiscence. Quantitative polymerase chain reaction analysis also indicated that *the Glyma09g06290* gene was highly expressed in pod indehiscent varieties [24].

Another gene regulating pod shattering in the domesticated soybean is *Pdh1*. *Pdh1* showed high homology to dirigent family genes, which were initially known as a stereoselective bimolecular phenoxy radical coupling of (E)-coniferyl alcohol, for producing lignan [25]. The functional *Pdh1* was found to be highly expressed in the lignified inner-sclerenchyma cells of the seed pod [26]. The inner sclerenchyma physical properties changed when *Pdh1* expression increased, and pod shattering started. As the relation between *Pdh1* and lignin was not clear yet, the gene might be responsible for lignin deposition in the seed pod. A non-synonymous nucleotide substitution that produces a stop codon results in pod-shattering-resistant varieties. Under low humidity *pdh1* allele containing soybeans showed significantly lower shattering scores than those with the *Pdh1* allele. This pod shattering-resistance associated allele was seen in more than 50% of Chinese and a considerable proportion of South Asian and North American landraces. However, Japanese and Korean landraces showed a very low frequency of this allele. The origin of domestication by selecting the indehiscent *Pdh1* allele might be originated from Huang-Huai-Hai Valley [22]. This infers that low humidity conditions provided selective pressure on *the pdh1* allele to protect seeds from dispersion. Zhang et al. (2022) provided a CRISPR/Cas9 gene editing solution for pod shuttering-susceptibility in a summer adapted soybean cultivar HC6 found in Huang-Huai-Hai [27]. They performed QTL mapping by using a recombinant inbred line population of HC6 and a pod shattering-resistant variety JD12 and they found a reproducible major allele at the *Pdh1* locus, SNP A/T that causes a nonsense variant (HC6/JD12). The resistant allele T was associated with low humidity regions in China, whereas the susceptible one A with high humidity regions in China, Japan, and Korea. Having known the contrasting effect, causal allele in different haplogroups facilitated the application of CRISPR/Cas9, the precise gene editing. This finally provided a gene therapy for pod shattering in soybean cultivars.

3.2 Shoot Growth Habit

Planting and harvesting time remarkably affect soybean yield, therefore, farmers must choose the appropriate maturity type regarding the environmental conditions. Soybean determinacy is an important agronomic trait that identifies the maturity type. Determinacy is governed by genes and environmental signals, which control

the generation of shoot apical meristem and transition to floral meristem. Soybean can be classified into three groups of determinacies: determinate, semi-determinate, and indeterminate. Indeterminate varieties, which are late maturing, show a prolonged vegetative phase with active stem and branch apices producing new nodes with leaves. Whereas determinate varieties, which are early maturing, cease stem and branch apical growth with photo-periodical floral induction.

Phenotypic variation amongst soybean landraces provided a good genetic resource for soybean breeding. Soybean planting management aims to maximize yield capacity and quality. It was found that, when indeterminate varieties are early-planted, they maintain an active vegetative growth for a long time and adequately accumulate amino acids and nutrients to allocate them towards seeds to increase yield quality and capacity. On the other hand, when determinate varieties are late-planted, yield capacity and moisture decrease. However, early planting of indeterminate soybean varieties can carry some risks. For example, late frost or extended pathogen infection might cause to decrease in yield capacity and even delay harvesting. To avoid the risks, determinate and indeterminate varieties are planted accordingly to maximize soybean production in the field [28–30].

In the cultivated soybean varieties, two genetic loci were identified to be associated with the determinacy trait: *Dt1* and *Dt2*. The *Dt1* allele is dominant or incompletely dominant on the *dt1* allele; the *Dt2* allele is dominant on the *dt2* allele. Soybean plants with *Dt1/Dt1* genotype are identified as indeterminate with *dt2/dt2* and semi-determinate with *Dt2/Dt2*. However, the *dt1/dt1* genotype shows a determinate phenotype when the *Dt2* locus is either recessive or dominant homozygous or heterozygous. Therefore, the *Dt1* locus has an epistatic effect on the *Dt2* locus [31, 32]. Their antagonistic behavior regulates flowering time and plant stem growth.

Dt1 is induced by *E3* and *E4* under long day conditions, interacts with bZIP family transcription factor FDc1, and binds to the promoter of *APETALA1* for delaying flowering. On the other way, when *APETALA1* binds to the promoter of *Dt1*, it inhibits its expression, thus promotes flowering [33]. *Dt1* locus encodes a phosphatidylethanolamine-binding protein (PEBP) family protein called GmTfl1 (or GmTfl1b) which is an ortholog of Arabidopsis TERMINAL FLOWER1 that controls plant height and internode length. *GmTfl1b* is expressed in the shoot apical meristem until flowering initiation [34]. Four independent single nucleotide polymorphisms on *GmTfl1* were identified in *Glycine max*, which makes non-synonymous amino acid changes, whereas these nucleotide changes were not observed in *Glycine soja.* This infers that the determinacy trait was introduced by the domestication process of soybean.

Four homologous *Tfl1* genes have been found in soybean. *Tfl1b* has been already known but the functions of other homologous genes are still not clear. Wang et al. (2023) identified the function of *Tfl1c* and *Tfl1d* by using CRISPR/Cas9 knock out double mutation [33]. Results showed that *tfl1c/tfl1d* double mutant soybeans flowered earlier than the wild type. The interaction of these homologous genes with *APETALA1* was also confirmed. Likewise, *Tfl1c* and *Tfl1d* interacted with FDc1 and inhibited four homologous genes of *APETALA1* and thus delaying flowering.

Dt2 encodes the MADS-domain factor that binds to the promoter of *Dt1* and inhibits the plant stem growth to start the transition from the vegetative phase to the

reproductive phase in soybean. There are two homologs of *the SUPPRESSOR OF OVEREXPRESSION OF CONSTANS1 (SOC1)* gene, known as the positive regulator of flowering which is activated downstream of *FLOWERING LOCUS T* (*FT*) in *Arabidopsis thaliana* [35], in soybean: *GmSOC1a* and *GmSOC1b.* These homologous genes interact in parallel with the *Dt2* locus and initiate flowering and inhibit stem elongation and node generation. A CRISPR/Cas9 gene knock out application mediated by *Agrobacterium tumefaciens* produced homozygous *soc1a, soc1b,* and *soc1a/soc1b* null mutants. Due to the lack of interaction between *Dt2* and *SOC1* in these mutant soybeans, *Dt1* expression was increased that caused delayed flowering.

Gibberellic acid biosynthesis is also very important for prolonged shoot growth. Plant height regulator genes in soybean, such as the dwarf gene *GmDW1, LATE ELONGATED HYPOCOTYL* (*GmLHY*), and an R2R3 MYB transcription factor *GmGAMYB* positively regulate gibberellic acid pathway [32]. A CRISPR/Cas9 knock out mutation application mediated by *Agrobacterium tumefaciens* produced a quadruple *LHY* mutant soybean (*Gmlhy1a,1b,2a,2b*), where it showed decreased plant height and gibberellic acid (GA3) level [19].

These functionally identified genes have provided insights into the molecular mechanisms underlying determinacy in soybean. Further studies utilizing techniques like CRISPR/Cas9 gene editing have deepened our understanding of these genes' functions and their interactions. Overall, unravelling the genetic and environmental factors controlling determinacy in soybean will contribute to the development of improved varieties with enhanced yield capacity and quality.

3.3 Photo-Periodicity

Soybean is a photo-period sensitive short-day plant, so it does not flower under long-day conditions. However, its adaptation to large latitudes requires a range of genetic variations of short-day activated genes. A group of genes has been identified to be involved in soybean latitudinal adaptation by regulating floral initiation.

Cryptochromes are generally known as blue-light receptor proteins involved in plant development and circadian clock. In *Arabidopsis thaliana* CRY1 was found to be responsible for mediation of blue-light induced de-etiolation, and CRY2 for photo-periodic flowering. *GmCRY1a* promotes blue-light-induced cell-wall elongation inhibition and, in contrast to *Arabidopsis thaliana,* regulates photo-periodic flowering by increasing the expression of *FT* mRNA [36].

The maturity loci in soybean, *E1, E2, E3,* and *E4* have a predominant effect on mediating photo-periodic flowering and maturity. The dominant *E1, E2, E3,* and *E4* allele delay flowering and the recessive allele facilitates adaptation to high latitudes by promoting early maturing, whereas *J, FT2a, FT5a, Time of flowering (Tof)16,* and *Tof18* facilitate adaptation to low latitudes [1, 37]. *J* locus, encoding a homolog of *EARLY FLOWERING 3 (ELF3)* in *Arabidopsis thaliana,* functions as a suppressor of the *E1* allele and promotes early flowering. On the other hand, *E1* suppresses *FT2a* and *FT5a,* and also, two homologs of the red-light photoreceptor phytochrome A (phyA) *E3* and *E4* interact with plant circadian clock evening complex (LUX)

and *E1* to delay flowering [37, 38]. Therefore, the phyA-LUX-E1-FT pathway regulates photo-periodic floral initiation in soybean. Cai et al. (2018) produced *GmFT2a* frameshift mutant soybeans having a 1 base-pair insertion or short deletion by applying sgRNAs CRISPR/Cas9 vectors through *Agrobacterium tumefaciens* mediated transformation [39]. They showed that the *ft2a* soybeans were flowering later than the wild types under short and long day photoperiodic conditions. Another study that targets the photoperiodicity of a short-day soybean variety showed that a frameshift mutation of *E1* caused truncated protein production. This unfunctional proteins thus disinhibited *GmFT2a/5a* and therefore, initiated flowering under long day conditions [40]. Another flowering repressor gene *GmPRR37* was identified by Wang et al. (2020) [41]. A CRISPR/Cas9 knock-out mutant *gmprr37* showed early flowering under long day conditions. Wang et al. (2023) clarified the function of *E2* and its homologous genes *E2-Like a* and *E2-Like b* [42]. They designed a single and double mutants of *E2* and its homologous genes through CRISPR/Cas9 knock out method to investigate their function in flowering and grain yield and their interaction with *E1*. It was found that *E2* is the major regulator of flowering but *E2-Like a* and *E2-Like b* were redundant. Only double mutants *e2/e2-like a* or *e2/e2-like b* initiated flowering earlier than *e2* types. Additionally, Li et al. (2022) discovered the interaction between photoperiodicity and plant protection by identifying the function of *GmCDPK38*, a calcium-dependent protein kinase encoding gene [43]. A cutworm susceptible soybean was mutated through CRISPR/Cas9 and *GmCDPK38* sgRNA vectors delivered by *Agrobacterium tumefaciens. gmcdpk38* soybeans delayed flowering and produced more defence-related metabolites under long-day conditions. This might suggest a beneficial soybean improvement strategy for growing resistant soybean at low latitudes. Recently, a new locus identified from GWAS, named *Tof8,* encodes a homolog of *the Arabidopsis thaliana FLAVIN-BINDING, KELCH REPEAT, F-BOX 1 (FKF1)* gene. There are two orthologous *FKF1* genes were found in soybean: *FKF1a* and *FKF1b.* These orthologous genes were found to be involved in flowering time and maturity by activating *E1* transcription through CRISPR/Cas9 knock-out mutation and two haplotypes of *FKF1b* were identified to be a part of regulating the latitudinal effect on flowering which might be a cause of natural selection during environmental adaptation [37].

Soybean flowering is dependent on a short photoperiod, so it usually flowers in spring or fall. Unravelling the mechanism underlying the latitudinal effect on flowering time will support soybean breeding across a wide range of latitudes. Moreover, it will be possible to produce long-day soybean varieties through genetic manipulation to speed up the breeding process.

3.4 Seed Quality

Seed traits determine the quality of the product. The seed coat, seed oil, oligosaccharides, lipoxigenases and protein contents are the important seed traits, which were improved during the domestication process. Breeding practices has shaped the modern soybean seed traits and improved the plant for a better human digestibility,

consumption, harvestability and endurance. Gene editing accelerates this process by functionally identifying the genes of interest associated with the desired seed traits.

Improved seed coat water impermeability protects embryos for sustainable seed storage and provides resistance to environmental deterioration in the field, such as pathogens, and mechanical and imbibition damages. Seed coat impermeability correlates with seed viability and longevity in soybean breeding programs. Additionally, hard seeds contain high calcium levels and increase the nutritional value of soyfoods [44]. As an adverse effect, hard seededness is unfavourable during post-harvest processing for vegetable oil and soyfood [45, 46]. Indeed, high seed water permeability facilitates water absorption and makes the seed easy to germinate. Jang et al. (2015) identified a seed hardness locus q*HS1*, which encodes an endo-1,4-β-glucanase [45]. A single nucleotide substitution from A to G in this gene dysfunctions the substrate-enzyme cleft domain and causes permeable seed coat in soybean. Likewise, a single nucleotide substitution from C to T by using *Agrobacterium tumefaciens* mediated transformations in *the GmHs1–1* gene, which encodes a calcineurin-like metallophosphoesterase transmembrane protein, showed increased permeability in the soybean seed coat [44]. Chandra et al. (2020) made inter-specific crosses between *Glycine max* and *Glycine soja* to understand the genetic inheritance of seed coat impermeability by using 217 recombinant inbred lines [46]. They phenotyped seed coat impermeability by looking into slow and rapid imbibition rates of the offspring. They identified three linked markers on chromosome 2, this locus was previously identified by Sun et al. (2015) and Jang et al. (2015) [44, 45]. Additionally, the phenotyping results revealed semi-permeable genotypes that might cause by minor alleles, and one of them was found to be associated with leaflet width, phytophthora resistance, and seed tocopherol. This implies that seed coat-identifying genes diversify in nature, and they maintain seed protection and coat-related alleles. The process of soybean breeding to improve seed coat impermeability should consider the involvement of the minor alleles as a potential genetic gain.

Another seed trait that was subjected to artificial selection during the domestication process is seed oil content. Cultivated soybean seeds contain more oil than wild seeds, which shows the effect of domestication on selecting high oil capacity in the seeds [47]. Soybean oil has a great economical value in the market, after palm oil, it is the second most-produced vegetable oil in the world between 2018 and 2023. China is the world leader in the production and consumption of soybean oil (USDA, 2023). It is used for human consumption and shows tremendous health benefits. Soybean oil consists of 15% saturated and 85% unsaturated fatty acids, which is responsible for lowered blood cholesterol levels, and decreased coronary heart disease [48]. On the other hand, soybean seeds became one of the most consumed vegan proteins for vegetarian and vegan diets. Seed oil increases with sugar mobilization towards the embryo. The underlying reason was investigated and found that sugar flux-controlling genes are upregulated and provide precursors for fatty acid biosynthesis. However, increasing seed oil content adversely affects seed protein content, which eventually influences the consumption of soybean as a source of protein.

Sugars Will Eventually be Exported Transporters (SWEET) gene family is responsible for the sugar transportation across cellular membranes having a role in the mobilization of carbohydrates from source to sink organs to sustain healthy growth and development of a plant [49, 50]. Bi-parental QTL mappings and genome-wide association studies dissect the genetic mechanisms by revealing QTLs on chromosomes 5, 7, 8, 10, 15, and 20, which are associated with seed oil, sugar, and protein contents [15, 51]. A major QTL has been identified on chromosome 15, *GmSWEET10a* (Glyma.15G049200), that affects seed oil, size, and protein content [15, 47]. This locus encodes a member of the SWEET gene family, a sugar transporter gene, ensuring sucrose efflux and allocating sugar from the mother seed coat to the filial embryo. Wang et al. (2020) showed that the frequency of the allele, which is significantly associated with seed oil content increase, is higher among landraces and cultivated soybeans than wild varieties [52]. Therefore, there was a strong artificial selection during the domestication of soybean. Zhang et al. (2020) identified a two-base-pair CC deletion in exon 6 of the cultivated and high-seed oil-containing soybean varieties [47]. They also unravelled that this gene shows a pleiotropic effect on seed oil and protein contents since the varieties having CC alleles available are significantly rich in protein. Additionally, Wang et al. (2020) identified a homologous locus of *GmSWEET10a*, named *GmSWEET10b*, by conducting a knock-out mutation through CRISPR/Cas9 showing a similar effect on seed oil and protein content while changing seed size; however, they could not find a significant artificial selection for this locus [52]. Cai et al. (2023) showed a similar antagonistic effect between seed oil and protein contents by indicating contrasting synthesis of oil and protein under changing expression of *the GmMFT* gene through CRISPR/Cas9 mediated knock-out mutants [48]. The reason underlying this antagonistic behavior might be that during the domestication process, carbohydrate transportation to seeds enhances fatty acid biosynthesis rather than protein synthesis. Wang et al. (2019) showed that the metabolic composition of soybean seeds unravelled how protein-rich seeds produced nitrogen-assimilating amino acids -free asparagine, free 3-cyanoalanine, free glutamic acid, L-malic acid, free glutamine, and free aspartic acid- and at the same time, they expressed the negative correlation to seed oil content [15]. Therefore, for a rich protein seed content synthesis of these amino acids are crucial.

Research on soybean seed oil-rich varieties always shows big seed in size. This might explain why seed size-regulating genes co-segregate with seed oil-regulating genes. However, seed size-regulating genes do not always correlate with seed oil or protein contents. Silencing of the *GmFAD3* gene, encoding omega-3 fatty acid desaturase, showed increased seed size without changing seed protein and oil content [53]. This implies that the seed size trait controlling *GmFAD3* is not linked with seed sugar efflux-controlling genes. Likewise, *the soybean seed size 1* (*GmSSS1*) gene, a homolog of *the Arabidopsis thaliana SPY* gene encodes an O-GlcNAc transferase protein and controls seed and pod size by showing pleiotropy, that was identified through CRISPR/Cas9 mediated knock-out mutants, and it might be achieved through cell expansion and division [54]. The independence of seed size-regulating

genes implies the epistatic effect of the size-regulating genes on *GmSWEET10a, GmSWEET10b,* and *GmMFT*.

Seed protein content is not the sole trait correlating with seed oil content, seed shape and coat color regulating genes are also co-segregating with the oil content. Soybean seed shape associated seed thickness (*ST1*) locus, encoding a UDP-D-glucuronate 4-epimerase, shows a pleiotropic effect on seed color and oil content through CRISPR/Cas9 mediated knock-out mutants. It acts upon seed shape by turning a flat seed into a round and catalyses the biosynthesis of UDP-galacturonic acid and participates in the glycolytic pathway. The conversion of UDP-glucuronic acid to UDP-galacturonic acid, which is a pectin precursor, promotes cell wall protein production. Simultaneously, glycerol-3-phosphate esterification produces triacylglycerol, which is a lipid biosynthesis precursor, and seeds show high oil content. The population analysis of 1209 soybean accessions, including 122 wild accessions, 542 landraces, and 545 cultivated soybeans, one haplotype, which shows a C to T polymorphism at nucleotide 203 of the *ST1* locus, was found the most frequent within the population and was significantly associated with high seed oil content and round shape. Another very interesting finding was that seed coat color determining locus *l*, a highly unstable transposon-induced locus, when reversed, co-segregates with *ST1*. This might explain the high frequency of yellow, round, and high oil involving cultivated soybean seeds [74].

The anthropogenic impact of oil-rich soybean domestication cannot be overseen. One reason could be that seeds rich in oil might be a good source of energy storage in the human body, so a simultaneous increase in seed oil and size might drive people to select oil-rich varieties. Another reason might be that carbon and nitrogen are primarily required for high seed protein content, so under limited nitrogen resources, oil accumulating large seeds might have been selected by humans. However, this hypothesis does not explain why seed oil accumulation still competes with protein accumulation under nitrogen-sufficient conditions. Soybean protein is a crucial supplement for vegetarian diet. Therefore, upregulating the production of nitrogen assimilating amino acids will enhance the soybean protein level. Gene editing will provide confirmation of the gene functions that are involved in soybean protein production.

3.5 Abiotic and Biotic Stress Resistance

Changing climatic conditions and pollution of air, soil, and water create devastating effects on soybean production. Abiotic and biotic stress conditions cause vulnerable soybean plants; indeed, their negative effects can be inherited and result in yield loss. Major abiotic stress conditions around the world for soybean are drought, salinity, cold, and flooding stresses [55]. For example, soybean production has been declining in Argentina, one of the major soybean producers, due to drought stress, which decreases the production, and so does the export and crushing [56]. QTL studies identified causal loci, which can promote marker-assisted breeding for

abiotic and biotic stress resistance. However, only a few of them have been validated through gene editing; otherwise, most of them are still in the candidate status.

The improved root system, enhanced water uptake, effective stomatal conductance, and slow wilting are some of the avoidance strategies in soybean from drought stress. Soybean breeding in light of genetic mechanisms can provide drought-resistant soybean crops and save soybean production under drought conditions. Throughout the investigations, several drought resistance-conferring genes have been functionally identified. A crosstalk between plant hormones and transcription factors regulates plant response to stress conditions. NAC (NAM, ATAF, and CUC) [57, 58], MYB [59], WRKY [59], AREB [60], DREB [61, 62], AP2/ERF [63] transcription factors were found to be involved in this collaboration. Overexpression of soybean *GmNFYA13,* a nuclear localization protein, was found to be responsible for gaining resilience to salt and drought stress in transgenic soybean plants. Abscisic acid (ABA) is one of the plant hormones which control the physiological adaptations of a plant under stress conditions. For example, stomatal closure is induced by increasing ABA to prevent water loss. When ABA is artificially induced in soybean, *GmNFYA13* expression was increased. This infers that *the GmNFYA13* gene is involved in abscisic acid-mediated stress response in soybean plants [64].

Soybean is a salt-sensitive plant, increased Na^+ ions change cellular ion balance and damage cells. *Cation Diffusion Facilitator 1, Arabidopsis* K^+ *Transporter 1,* and also some transcription factors that are generally involved in abiotic and biotic stress conditions, such as MYB, WRKY, AP2/ERF, and NAC are associated with salt stress resistance in soybean [65]. Wang et al. (2021) identified an ABA and salt induced transcription repressor *GmAITR* in soybean to reduce the salinity stress related phenotypes without losing its fitness [66]. A CRISPR/Cas9 knock out mutant technique, *gmaitr* inhibited the expression of ABA and showed tolerance to salt stress.

Moreover, flooding stress is another abiotic stress that affects soybean production under ill-drained soils. Soybean roots are primary organs that are affected by flooding stress, limited oxygen uptake causes hypoxia and reduced energy production. To overcome this stress, plants undergo alternative energy-producing metabolic activities. Transcriptomic and proteomic studies unravel a group of proteins that are involved in cell wall modification, methylglyoxal detoxification, hypoxia reduction, pathogen defence, reactive oxygen species scavengers and chaperons, and energy production through glycolysis induction and alcohol fermentation [67–69].

Soybean production around the world is challenged by increasing negative impacts of fungus, bacterium, phytoplasma, nematode, and virus infections. Natural and artificial selection strategies improved soybean resistance over the years and sustain its development and reproduction despite dynamic spatial and temporal conditions. Soybean breeding for biotic stress resistance is a very active process. Due to changing climate conditions, pathogen populations shift, and new races are introduced to host plants. This activates new protection mechanisms and beneficial mutations in resistance-conferring genes provide endurance to plants. The selection pressure on beneficial mutations can occur both naturally and artificially. Zhao et al.

(2015) investigated nucleotide fixation of pathogen resistance in wild and cultivated varieties, and their study revealed that *Glyma20g08290* (homolog of *Arabidopsis thaliana RPM1* gene) is a naturally selected locus, which is associated with *Pseudomonas syringae* in soybean (Ashfield et al., 1995) and found in wild soybean varieties [70, 71].

Marker-assisted selection provides causal QTLs for vertical and horizontal resistance. It unravels major R genes, which maintain vertical resistance in soybean to soybean cyst nematode (*Rhg*), Phytophthora root and stem rot (*Rps*), soybean rust (*Rpp*), frog eye leaf spot (*Rcs*), bacterial blight (*Rpg*), and soybean mosaic virus (*Rsv* and *Rsc*) [65, 70, 72, 73]. R genes provide full protection in a race-specific manner. Horizontal resistance is controlled by multiple minor effect genes and confers resistance against many soybean diseases such as sudden death syndrome, Sclerotinia stems rot, root-knot nematode, and most Pythium species [73]. However, this type of protection is not pathotype specific so it is more long-lasting than vertical resistance. In vertical resistance, environmental conditions might cause genetic changes in avirulent proteins, which are recognized by pathotype-specific R genes, or shift in pathogen populations. On the other hand, utilization of molecular markers associated with R genes is more feasible than pursuing a soybean breeding strategy for minor allelic resistance.

4 Conclusion

QTLs have been identified for a number of traits in soybean. These QTLs can be used to develop marker-assisted breeding programs to improve soybean cultivars for resistance to these stresses. Gene editing is a newer technology that can be used to rapidly and efficiently introduce and edit specific genes. Gene editing is a complementary approach to marker-assisted breeding, and the two technologies can be used together to accelerate the development of improved soybean cultivars. The advantages of using gene editing for soybean improvement:

- Gene editing is a precise technology that can be used to target specific genes.
- Gene editing is a rapid technology that can be used to develop new cultivars in a shorter time frame than conventional breeding methods.

The challenges of using gene editing for soybean improvement:

- Gene editing is a regulated technology, and there are a number of regulatory hurdles that must be overcome before gene-edited soybeans can be commercialized.
- There is some public opposition to the use of gene editing in food crops.

Despite the challenges, gene editing is a promising technology that has the potential to revolutionize soybean improvement. By combining gene editing with marker-assisted breeding, we can develop soybean cultivars that are more resistant to

abiotic and biotic stresses, have improved yield and nutritional quality, and are better suited to the changing climate.

References

1. Sedivy, E.J., Wu, F., Hanzawa, Y.: Soybean domestication: the origin, genetic architecture and molecular bases. New Phytol. **214**, 539–553 (2017)
2. Ritchie, H., Roser, M.: Forests and deforestation. Our World in Data. Retrieved from https://ourworldindata.org/forests-and-deforestation (2021)
3. Fraanje, W., Garnett, T.: Soy: Food, Feed, and Land Use Change. (Foodsource: Building Blocks). Food Climate Research Network. University of Oxford (2020)
4. Baloglu, M.C., Celik Altunoglu, Y., Baloglu, P., Yildiz, A.B., Türkölmez, N., Özden Çiftçi, Y.: Gene-editing technologies and applications in legumes: Progress, evolution, and future prospects. Front. Genet. **13**, 859437 (2022)
5. Curtin, S.J., Zhang, F., Sander, J.D., Haun, W.J., Starker, C., Baltes, N.J., Reyon, D., Dahlborg, E.J., Goodwin, M.J., Coffman, A.P., Dobbs, D., Joung, J.K., Voytas, D.F., Stupar, R.M.: Targeted mutagenesis of duplicated genes in soybean with zinc-finger nucleases. Plant Physiol. **156**, 466–473 (2011)
6. Curtin, S.J., Michno, J.M., Campbell, B.W., Gil-Humanes, J., Mathioni, S.M., Hammond, R., Gutierrez-Gonzalez, J.J., Donohue, R.C., Kantar, M.B., Eamens, A.L., Meyers, B.C., Voytas, D.F., Stupar, R.M.: MicroRNA maturation and MicroRNA target gene expression regulation are severely disrupted in soybean dicer-like1 double mutants. G3 (Bethesda, Md.). **6**, 423–433 (2016)
7. Bonawitz, N.D., Ainley, W.M., Itaya, A., Chennareddy, S.R., Cicak, T., Effinger, K., Jiang, K., Mall, T.K., Marri, P.R., Samuel, J.P., Sardesai, N., Simpson, M., Folkerts, O., Sarria, R., Webb, S.R., Gonzalez, D.O., Simmonds, D.H., Pareddy, D.R.: Zinc finger nuclease-mediated targeting of multiple transgenes to an endogenous soybean genomic locus via non-homologous end joining. Plant Biotechnol. J. **17**, 750–761 (2019)
8. Haun, W., Coffman, A., Clasen, B.M., Demorest, Z.L., Lowy, A., Ray, E., Retterath, A., Stoddard, T., Juillerat, A., Cedrone, F., Mathis, L., Voytas, D.F., Zhang, F.: Improved soybean oil quality by targeted mutagenesis of the fatty acid desaturase 2 gene family. Plant Biotechnol. J. **12**, 934–940 (2014)
9. Demorest, Z.L., Coffman, A., Baltes, N.J., Stoddard, T.J., Clasen, B.M., Luo, S., Retterath, A., Yabandith, A., Gamo, M.E., Bissen, J., Mathis, L., Voytas, D.F., Zhang, F.: Direct stacking of sequence-specific nuclease-induced mutations to produce high oleic and low linolenic soybean oil. BMC Plant Biol. **16**, 225 (2016)
10. Du, H., Zeng, X., Zhao, M., Cui, X., Wang, Q., Yang, H., Cheng, H., Yu, D.: Efficient targeted mutagenesis in soybean by TALENs and CRISPR/Cas9. J. Biotechnol. **217**, 90–97 (2016)
11. Chen, X., Yang, S., Zhang, Y., Zhu, X., Yang, X., Zhang, C., Li, H., Feng, X.: Generation of male-sterile soybean lines with the CRISPR/Cas9 system. Crop J. **9**, 1270–1277 (2021)
12. Li, Z., Cheng, Q., Gan, Z., Hou, Z., Zhang, Y., Li, Y., Li, H., Nan, H., Yang, C., Chen, L., Lu, S., Shi, W., Chen, L., Wang, Y., Fang, C., Kong, L., Su, T., Li, S., Kou, K., Wang, L., Kong, F., Liu, B., Dong, L.: Multiplex CRISPR/Cas9-mediated knockout of soybean LNK2 advances flowering time. Crop J. **9**, 767–776 (2021)
13. Xiao, Z., Jin, Y., Zhang, Q., Lamboro, A., Dong, B., Yang, Z., Wang, P.: Construction and functional analysis of CRISPR/Cas9 vector of FAD2 gene family in soybean. Phyton-Int. J. Exp. Bot. **91**, 349–361 (2022)
14. Zhou, J., Li, Z., Li, Y., Zhao, Q., Luan, X., Wang, L., Liu, Y., Liu, H., Zhang, J., Yao, D.: Effects of different gene editing modes of CRISPR/Cas9 on soybean fatty acid anabolic metabolism based on GmFAD2 family. Int. J. Mol. Sci. **24**, 4769 (2023)

15. Wang, J., Zhou, P., Shi, X., Yang, N., Yan, L., Zhao, Q., Yang, C., Guan, Y.: Primary metabolite contents are correlated with seed protein and oil traits in near-isogenic lines of soybean. Crop J. **7**, 651–659 (2019)
16. Cao, L., Wang, Z., Ma, H., Liu, T., Ji, J., Duan, K.: Multiplex CRISPR/Cas9-mediated raffinose synthase gene editing reduces raffinose family oligosaccharides in soybean. Front. Plant Sci. **13**, 1048967 (2022)
17. Lin, W., Kuang, H., Bai, M., Jiang, X., Zhou, P., Li, Y., Chen, B., Li, H., Guan, Y.: Multiplex genome editing targeting soybean with ultra-low anti-nutritive oligosaccharides. Crop J.
18. Bao, A., Chen, H., Chen, L., Chen, S., Hao, Q., Guo, W., Qiu, D., Shan, Z., Yang, Z., Yuan, S., Zhang, C., Zhang, X., Liu, B., Kong, F., Li, X., Zhou, X., Tran, L.P., Cao, D.: CRISPR/Cas9-mediated targeted mutagenesis of GmSPL9 genes alters plant architecture in soybean. BMC Plant Biol. **19**, 131 (2019)
19. Cheng, Q., Dong, L., Su, T., Li, T., Gan, Z., Nan, H., Lu, S., Fang, C., Kong, L., Li, H., Hou, Z., Kou, K., Tang, Y., Lin, X., Zhao, X., Chen, L., Liu, B., Kong, F.: CRISPR/Cas9-mediated targeted mutagenesis of GmLHY genes alters plant height and internode length in soybean. BMC Plant Biol. **19**, 562 (2019)
20. Spengler, R.N.: Anthropogenic seed dispersal: rethinking the origins of plant domestication. Trends Plant Sci. **25**, 340–348 (2020)
21. Dong, Y., Yang, X., Liu, J., et al.: Pod shattering resistance associated with domestication is mediated by a NAC gene in soybean. Nat. Commun. **5** (2014)
22. Zhang, J., Singh, A.K.: Genetic control and geo-climate adaptation of pod dehiscence provide novel insights into soybean domestication. G3 (Bethesda, Md.). **10**, 545–554 (2020)
23. Parker, T.A., Lo, S., Gepts, P.: Pod shattering in grain legumes: emerging genetic and environment-related patterns. Plant Cell. **33**, 179–199 (2021)
24. Hu, D., Kan, G., Hu, W., Li, Y., Hao, D., Li, X., Yang, H., Yang, Z., He, X., Huang, F., et al.: Identification of loci and candidate genes responsible for pod dehiscence in soybean via genome-wide association analysis across multiple environments. Front. Plant Sci. **10**, 811 (2019)
25. Davin, L.B., Wang, H.-B., Crowell, A.L., Bedgar, D.L., Martin, D.M., Sarkanen, S., Lewis, N.G.: Stereoselective bimolecular Phenoxy radical coupling by an auxiliary (Dirigent) protein without an active center. Science. **275**, 362–366 (1997)
26. Funatsuki, H., Suzuki, M., Hirose, A., Inaba, H., Yamada, T., Hajika, M., Komatsu, K., Katayama, T., Sayama, T., Ishimoto, M., et al.: Molecular basis of a shattering resistance boosting global dissemination of soybean. Proc. Natl Acad. Sci. USA. **111**, 17797–17802 (2014)
27. Zhang, Z., Wang, J., Kuang, H., et al.: Elimination of an unfavorable allele conferring pod shattering in an elite soybean cultivar by CRISPR/Cas9. aBIOTECH. **3**, 110–114 (2022)
28. Singh, M., Siler, T. Soybean planting considerations for maximum profits. MSU Extension Soybeans. https://www.canr.msu.edu/news/soybean-planting-and-time-management-considerations (2022)
29. Staton, M. Reducing soybean harvest losses in 2017. MSU Extension Soybeans. https://www.canr.msu.edu/news/reducing_soybean_harvest_losses_in_2017 (2017)
30. Staton, M. Overview of 2020 soybean on-farm research projects. MSU Extension Soybeans. https://www.canr.msu.edu/news/soybean_on_farm_research_trials (2020)
31. Tian, X., Wang, X., Lee, R., Li, Y., Specht, J.E., Nelson, R.L., McClean, P.E., Qiu, L., Ma, J.: Artificial selection for determinate growth habit in soybean. Proc. Natl. Acad. Sci. **107**, 8563–8568 (2010)
32. Zhou, X., Wang, D., Mao, Y., Zhou, Y., Zhao, L., Zhang, C., Liu, Y., Chen, J.: The organ size and morphological change during the domestication process of soybean. Front. Plant Sci. **13**, 913238 (2022)
33. Wang, L., Lin, C., Li, B., Su, T., Li, S., Li, H., He, F., Gou, C., Chen, Z., Wang, Y., Qin, J., Liu, B., Kong, F., Yue, L., Lu, S., Fang, C.: Two soybean homologues of TERMINAL FLOWER 1 control flowering time under long day conditions. Crop J. **11**(3), 704–712 (2023)

34. Liu, B., Watanabe, S., Uchiyama, T., Kong, F., Kanazawa, A., Xia, Z., Nagamatsu, A., Arai, M., Yamada, T., Kitamura, K., Masuta, C., Harada, K., Abe, J.: The soybean stem growth habit gene Dt1 is an Ortholog of Arabidopsis TERMINAL FLOWER1. Plant Physiol. **153**, 198–210 (2010)
35. Yoo, S.K., Chung, K.S., Kim, J., Lee, J.H., Hong, S.M., Yoo, S.J., Yoo, S.Y., Lee, J.S., Ahn, J.H.: CONSTANS activates SUPPRESSOR OF OVEREXPRESSION OF CONSTANS 1 through FLOWERING LOCUS T to promote flowering in Arabidopsis. Plant Physiol. **139**, 770–778 (2005)
36. Zhang, Q., Li, H., Li, R., Hu, R., Fan, C., Chen, F., Wang, Z., Liu, X., Fu, Y., Lin, C.: Association of the circadian rhythmic expression of GmCRY1a with a latitudinal cline in photoperiodic flowering of soybean. Proc. Natl. Acad. Sci. U. S. A. **105**, 21028–21033 (2008)
37. Li, H., Du, H., He, M., Wang, J., Wang, F., Yuan, W., Huang, Z., Cheng, Q., Gou, C., Chen, Z., Liu, B., Kong, F., Fang, C., Zhao, X., & Yu, D.: Natural variation of FKF1 controls flowering and adaptation during soybean domestication and improvement. New Phytol. Advance online publication. (2023)
38. Zhao, F., Lyu, X., Lyu, X., Ji, R., Ji, R., Liu, J., Liu, J., Zhao, T., Zhao, T., Li, H., Li, H., Liu, B., Liu, B., Pei, Y., Pei, Y.: CRISPR/Cas9-engineered mutation to identify the roles of phytochromes in regulating photomorphogenesis and flowering time in soybean. Crop J. **10**, 1654–1664 (2022)
39. Cai, Y., Chen, L., Liu, X., Guo, C., Sun, S., Wu, C., Jiang, B., Han, T., Hou, W.: CRISPR/Cas9-mediated targeted mutagenesis of GmFT2a delays flowering time in soya bean. Plant Biotechnol. J. **16**, 176–185 (2018)
40. Han, J., Guo, B., Guo, Y., Zhang, B., Wang, X., Qiu, L.J.: Creation of early flowering germplasm of soybean by CRISPR/Cas9 technology. Front. Plant Sci. **10**, 1446 (2019)
41. Wang, L., Sun, S., Wu, T., Liu, L., Sun, X., Cai, Y., Li, J., Jia, H., Yuan, S., Chen, L., Jiang, B., Wu, C., Hou, W., Han, T.: Natural variation and CRISPR/Cas9-mediated mutation in GmPRR37 affect photoperiodic flowering and contribute to regional adaptation of soybean. Plant Biotechnol. J. **18**, 1869–1881 (2020)
42. Wang, L., Li, H., He, M., Dong, L., Huang, Z., Chen, L., Nan, H., Kong, F., Liu, B., Zhao, X.: GIGANTEA orthologs, E2 members, redundantly determine photoperiodic flowering and yield in soybean. J. Integr. Plant Biol. **65**, 188–202 (2023)
43. Li, X., Hu, D., Cai, L., Wang, H., Liu, X., Du, H., Yang, Z., Zhang, H., Hu, Z., Huang, F., Kan, G., Kong, F., Liu, B., Yu, D., Wang, H.: CALCIUM-DEPENDENT PROTEIN KINASE38 regulates flowering time and common cutworm resistance in soybean. Plant Physiol. **190**, 480–499 (2022)
44. Sun, L., Miao, Z., Cai, C., Zhang, D., Zhao, M., Wu, Y., Zhang, X., Swarm, S.A., Zhou, L., Zhang, Z.J., Nelson, R.L., Ma, J.: GmHs1-1, encoding a calcineurin-like protein, controls hard-seededness in soybean. Nat. Genet. **47**, 939–943 (2015)
45. Jang, S.J., Sato, M., Sato, K., Jitsuyama, Y., Fujino, K., et al.: A single-nucleotide polymorphism in an Endo-1,4-β-Glucanase gene controls seed coat permeability in soybean. PLoS One. **10**, e0128527 (2015)
46. Chandra, S., Taak, Y., Rathod, D.R., Yadav, R.R., Poonia, S., Sreenivasa, V., Talukdar, A.: Genetics and mapping of seed coat impermeability in soybean using interspecific populations. Physiol. Mol. Biol. Plants. **26**, 2291–2299 (2020)
47. Zhang, H., Goettel, W., Song, Q., Jiang, H., Hu, Z., Wang, M.L., An, Y.C.: Selection of GmSWEET39 for oil and protein improvement in soybean. PLoS Genet. **16**(11), e1009114 (2020). https://doi.org/10.1371/journal.pgen.1009114
48. Cai, Z., Xian, P., Cheng, Y., Zhong, Y., Yang, Y., Zhou, Q., Lian, T., Ma, Q., Nian, H., Ge, L.: MOTHER-OF-FT-AND-TFL1 regulates the seed oil and protein content in soybean. New Phytol. Advance online publication. (2023)
49. Chen, L.-Q., Hou, B.-H., Lalonde, S., Frommer, W.B.: Sugar transporters for intercellular exchange and nutrition of pathogens. Nature. **468**, 527–532 (2010)

50. Patil, G., Valliyodan, B., Deshmukh, R., Prince, S., Nicander, B., Zhao, M., Sonah, H., Song, L., Lin, L., Chaudhary, J., Liu, Y., Joshi, T., Xu, D., Nguyen, H.T., Mutava, R.N.: Soybean (Glycine max) SWEET gene family: insights through comparative genomics, transcriptome profiling and whole genome re-sequence analysis. BMC Genomics. **16**, 520 (2015)
51. Huang, J., Ma, Q., Cai, Z., Xia, Q., Li, S., Jia, J., Chu, L., Lian, T., Nian, H., Cheng, Y.: Identification and mapping of stable QTLs for seed oil and protein content in soybean [Glycine max (L.) Merr.]. J. Agric. Food Chem. **68**, 6448–6460 (2020)
52. Wang, S., Liu, S., Wang, J., Yokosho, K., Zhou, B., Yu, Y.C., Liu, Z., Frommer, W.B., Ma, J.F., Chen, L.Q., Guan, Y., Shou, H., Tian, Z.: Simultaneous changes in seed size, oil content and protein content driven by selection of SWEET homologues during soybean domestication. Natl. Sci. Rev. **7**, 1776–1786 (2020)
53. Singh, A.K., Fu, D.Q., El-Habbak, M., Navarre, D., Ghabrial, S., Kachroo, A.: Silencing genes encoding omega-3 fatty acid desaturase alters seed size and accumulation of bean pod mottle virus in soybean. Mol. Plant-Microbe Interact. **24**, 506–515 (2011)
54. Zhu, W., Yang, C., Yong, B., Wang, Y., Li, B., Gu, Y., Wei, S., An, Z., Sun, W., Qiu, L., He, C.: An enhancing effect attributed to a nonsynonymous mutation in SOYBEAN SEED SIZE 1, a SPINDLY-like gene, is exploited in soybean domestication and improvement. New Phytol. **236**, 1375–1392 (2022)
55. Chugh, V., Kaur, D., Purwar, S., Kaushik, P., Sharma, V., Kumar, H., Rai, A., Singh, C.M., Kamaluddin, Dubey, R.B.: Applications of molecular markers for developing abiotic-stress-resilient oilseed crops. Life (Basel, Switzerland). **13**, 88 (2022)
56. Boroughs, B.: Oilseeds and products annual (Report no. AR2023–0004). US department of agriculture: foreign agricultural service. https://apps.fas.usda.gov/newgainapi/api/Report/DownloadReportByFileName?fileName=Oilseeds%20and%20Products%20Annual_Buenos%20Aires_Argentina_AR2023-0004.pdf (2023)
57. Lu, P.L., Chen, N.Z., An, R., Su, Z., Qi, B.S., Ren, F., Chen, J., Wang, X.C., Tao, J.J.: A novel drought-inducible gene, ATAF1, encodes a NAC family protein that negatively regulates the expression of stress-responsive genes in Arabidopsis. Plant Mol. Biol. **63**, 289–305 (2007)
58. Yang, C., Huang, Y., Lv, P., Antwi-Boasiako, A., Begum, N., Zhao, T., Zhao, J.: NAC transcription factor GmNAC12 improved drought stress tolerance in soybean. Int. J. Mol. Sci. **23**, 12029 (2022)
59. Arya, H., Singh, M.B., Bhalla, P.L.: Towards developing drought-smart soybeans. Front. Plant Sci. **12**, 750664 (2021)
60. Fuganti-Pagliarini, R., Ferreira, L.C., Rodrigues, F.A., Molinari, H.B.C., Marin, S.R.R., Molinari, M.D.C., Marcolino-Gomes, J., Mertz-Henning, L.M., Farias, J.R.B., de Oliveira, M.C.N., Neumaier, N., Kanamori, N., Fujita, Y., Mizoi, J., Nakashima, K., Yamaguchi-Shinozaki, K., Nepomuceno, A.L.: Characterization of soybean genetically modified for drought tolerance in field conditions. Front. Plant Sci. **8**, 448 (2017)
61. Hou, Z., Li, Y., Cheng, Y., Li, W., Li, T., Du, H., Kong, F., Dong, L., Zheng, D., Feng, N., Liu, B., Cheng, Q.: Genome-wide analysis of DREB genes identifies a novel salt tolerance gene in wild soybean (Glycine soja). Front. Plant Sci. **13**, 821647 (2022)
62. Kidokoro, S., Watanabe, K., Ohori, T., Moriwaki, T., Maruyama, K., Mizoi, J., Myint Phyu Sin Htwe, N., Fujita, Y., Sekita, S., Shinozaki, K., Yamaguchi-Shinozaki, K.: Soybean DREB1/CBF-type transcription factors function in heat and drought as well as cold stress-responsive gene expression. Plant J. **81**, 505–518 (2015)
63. Zhang, G., Chen, M., Li, L., Xu, Z., Chen, X., Guo, J., Ma, Y.: Overexpression of the soybean GmERF3 gene, an AP2/ERF type transcription factor for increased tolerances to salt, drought, and diseases in transgenic tobacco. J. Exp. Bot. **60**, 3781–3796 (2009)
64. Ma, X.J., Fu, J.D., Tang, Y.M., Yu, T.F., Yin, Z.G., Chen, J., Zhou, Y.B., Chen, M., Xu, Z.S., Ma, Y.Z.: GmNFYA13 improves salt and drought tolerance in transgenic soybean plants. Front. Plant Sci. **11**, 587244 (2020)
65. Du, H., Fang, C., Li, Y., Kong, F., Liu, B.: Understandings and future challenges in soybean functional genomics and molecular breeding. J. Integr. Plant Biol. **65**, 468–495 (2023)

66. Wang, T., Xun, H., Wang, W., Ding, X., Tian, H., Hussain, S., Dong, Q., Li, Y., Cheng, Y., Wang, C., Lin, R., Li, G., Qian, X., Pang, J., Feng, X., Dong, Y., Liu, B., Wang, S.: Mutation of GmAITR genes by CRISPR/Cas9 genome editing results in enhanced salinity stress tolerance in soybean. Front. Plant Sci. **12**, 779598 (2021)
67. Komatsu, S., Yamamoto, R., Nanjo, Y., Mikami, Y., Yunokawa, H., Sakata, K.: A comprehensive analysis of the soybean genes and proteins expressed under flooding stress using transcriptome and proteome techniques. J. Proteome Res. **8**, 4766–4778 (2009)
68. Yin, X., Komatsu, S.: Comprehensive analysis of response and tolerant mechanisms in early-stage soybean at initial-flooding stress. J. Proteome. **169**, 225–232 (2017)
69. Khan, M.N., Ahmed, I., Ud Din, I., Noureldeen, A., Darwish, H., Khan, M.: Proteomic insight into soybean response to flooding stress reveals changes in energy metabolism and cell wall modifications. PLoS One. **17**, e0264453 (2022)
70. Zhao, S., Zheng, F., He, W., Wu, H., Pan, S., Lam, H.M.: Impacts of nucleotide fixation during soybean domestication and improvement. BMC Plant Biol. **15**, 81 (2015)
71. Ashfield, T., Keen, N.T., Buzzell, R.I., Innes, R.W.: Soybean resistance genes specific for different pseudomonas syringae avirulence genes are allelic, or closely linked, at the RPG1 locus. Genetics. **141**, 1597–1604 (1995)
72. Cook, D.E., Lee, T.G., Guo, X., Melito, S., Wang, K., Bayless, A.M., Wang, J., Hughes, T.J., Willis, D.K., Clemente, T.E., Diers, B.W., Jiang, J., Hudson, M.E., Bent, A.F.: Copy number variation of multiple genes at Rhg1 mediates nematode resistance in soybean. Science (New York, N.Y.). **338**, 1206–1209 (2012)
73. Lin, F., Chhapekar, S.S., Vieira, C.C., Da Silva, M.P., Rojas, A., Lee, D., Liu, N., Pardo, E.M., Lee, Y.C., Dong, Z., Pinheiro, J.B., Ploper, L.D., Rupe, J., Chen, P., Wang, D., Nguyen, H.T.: Breeding for disease resistance in soybean: a global perspective. TAG. Theor. Appl. Genet. Theoretische und angewandte Genetik. **135**, 3773–3872 (2022)
74. Li, J., Zhang, Y., Ma, R., Huang, W., Hou, J., Fang, C., Wang, L., Yuan, Z., Sun, Q., Dong, X., Hou, Y., Wang, Y., Kong, F., Sun, L.: Identification of ST1 reveals a selection involving hitchhiking of seed morphology and oil content during soybean domestication. Plant Biotechnol. J. **20**, 1110–1121 (2022)